A gene pathway analysis highlights the role of cellular adhesion molecules in multiple sclerosis susceptibility

Damotte, V ; Guillot-Noel, L ; Patsopoulos, N A ; Madireddy, L ; El Behi, M ; De Jager, P L ; Baranzini, S E ; Cournu-Rebeix, I ; Fontaine, B ; Martin, R

Abstract: Genome-wide association studies (GWASs) perform per-SNP association tests to identify variants involved in disease or trait susceptibility. However, such an approach is not powerful enough to unravel genes that are not individually contributing to the disease/trait, but that may have a role in interaction with other genes as a group. Pathway analysis is an alternative way to highlight such group of genes. Using SNP association P-values from eight multiple sclerosis (MS) GWAS data sets, we performed a candidate pathway analysis for MS susceptibility by considering genes interacting in the cell adhesion molecule (CAMs) biological pathway using Cytoscape software. This network is a strong candidate, as it is involved in the crossing of the blood-brain barrier by the T cells, an early event in MS pathophysiology, and is used as an efficient therapeutic target. We drew up a list of 76 genes belonging to the CAM network. We highlighted 64 networks enriched with CAM genes with low P-values. Filtering by a percentage of CAM genes up to 50% and rejecting enriched signals mainly driven by transcription factors, we highlighted five networks associated with MS susceptibility. One of them, constituted of ITGAL, ICAM1 and ICAM3 genes, could be of interest to develop novel therapeutic targets.

DOI: https://doi.org/10.1038/gene.2013.70

A gene pathway analysis highlights the role of cellular adhesion molecules in Multiple Sclerosis susceptibility

Vincent Damotte[1], Léna Guillot-Noel[1], Nikolaos A. Patsopoulos[2,3,4,5], Lohith Madireddy[6], Mohamed El Behi[1,7], International Multiple Sclerosis Genetics Consortium[¶], Wellcome Trust Case Control Consortium 2[¶], Philip L. De Jager[2,4,5], Sergio E. Baranzini[6], Isabelle Cournu-Rebeix[1,8]*, Bertrand Fontaine[1,8]*

* equal contribution

[1]CRICM-UPMC-INSERM-UMR_S 975-CNRS UMR7225 75013 PARIS France

[2]Program in Translational NeuroPsychiatric Genomics, Institute for the Neurosciences, Departments of Neurology and Psychiatry, Brigham & Women's Hospital, Boston, 02115 MA, USA

[3]Division of Genetics, Department of Medicine, Brigham & Women's Hospital, Harvard Medical School, Boston, MA, USA

[4]Harvard Medical School, Boston, MA 02115, USA

[5]Broad Institute of Harvard and Massachusetts Institute of Technology, Cambridge, MA, USA

[6]Department of Neurology, University of California at San Francisco, Sandler Neurosciences Center, 675 Nelson Rising Lane, San Francisco, CA 94158, USA

[7]Institut du Cerveau et de la Moelle épinière, ICM, 75013 Paris, France

[8]Département de Neurologie Pitié-Salpêtrière AP-HP 75013 PARIS France

[¶] Membership of the International Multiple Sclerosis Genetics Consortium (IMSGC) and the Wellcome Trust Case Control Consortium 2 (WTCCC2) is provided in the Acknowledgments

Corresponding author :

Bertrand FONTAINE

Hôpital Pitié-Salpêtrière - Bâtiment ICM

UMR_S975 - Equipe Bertrand Fontaine - 4ème étage

47 boulevard de l'Hôpital

75013 PARIS, France

T : 01 57 27 44 10 / F : 01 57 27 47 92 / e-mail : bertrand.fontaine@upmc.fr

Abstract:

Typically Genome-Wide Association Studies (GWASs) perform per-SNP association tests to identify variants involved in disease or trait susceptibility. However, such an approach is not powerful enough to unravel genes that are not individually contributing to the disease/trait but that may play a role in interaction with other genes as a group. Pathway analysis is an alternative way to highlight such group of genes. Using SNPs association p-values from 8 Multiple Sclerosis (MS) GWAS datasets, we performed a candidate pathway analysis for MS susceptibility considering genes interacting in cell adhesion molecules (CAMs) biological pathway using Cytoscape software. This network is a strong candidate since it is involved in the crossing of the blood brain barrier by the T cells, an early event in MS pathophysiology, and used as an efficient therapeutic target. We drew up a list of 76 genes belonging to the CAMs network. We highlighted 64 networks enriched with CAMs genes with low p-values. Filtering by a percentage of CAMs genes up to 50% and rejecting enriched signals mainly driven by transcription factors, we highlighted 5 networks associated with MS susceptibility. One of them, constituted of *ITGAL*, *ICAM1* and *ICAM3* genes could be of interest to develop novel therapeutic targets.

Keywords: Multiple Sclerosis, Susceptibility, Pathway Analysis, Cell Adhesion Molecules

Introduction

Multiple Sclerosis (MS) is a common inflammatory and demyelinating disease of the central nervous system [1]. Epidemiological studies have proved the multifactorial causes of the disease, resulting from the interaction between genetic factors and currently unknown environmental factors [2].

To date, MHC genetic variants, as well as 110 non-MHC variants, have been associated with MS susceptibility by GWAS[3-7]. However, it has been estimated that several other genes contribute to disease susceptibility and are yet to be identified[8].

Several hypotheses have been proposed to explain the missing heritability[9,10]. One of them points at interactions between genes involved in the same biological pathway[11, 12]. According to this hypothesis, some of the genes implicated in MS susceptibility cannot, on their own, reach GWAS significance but grouped in a pathway may collectively contribute to MS genetic component.

The initial event leading to the development of an MS lesion is blood-brain-barrier (BBB) disruption and the crossing of the latter by peripherally activated T cells [13]. This step, which is a key mechanism in our current understanding of MS physiopathology, requires the interaction of integrins, cell surface molecules expressed by T cells, to the adhesion molecules expressed by the BBB endothelium, ultimately allowing the lymphocytes transmigration into the brain[14-17].

The importance of the BBB crossing in MS physiopathology is stressed by the use of treatments active on adhesion molecules. For example, Natalizumab, a monoclonal antibody blocking the interaction between VLA-4 (very late antigen-4) molecule (composed of ITGB1, integrin beta 1, and ITGA4, integrin alpha 4, integrins subunits) and its receptor VCAM1 (vascular cell adhesion molecule-1), is among the most efficient therapies available for MS patients to date [18, 19]. The role of adhesion molecules in MS has also been reinforced by the demonstration that the expression of ALCAM (activated leucocyte cell adhesion molecule) located at the surface of the BBB endothelium, was increased during the inflammatory process in MS patients. In the same study, it has been shown that in vitro expression of ALCAM, ICAM-1 (intercellular cell adhesion molecule-1) and VCAM-1 was dependent of activation of the human BBB endothelial cells by pro-inflammatory cytokines (TNF, Tumor Necrosis Factor et IFN-γ, Interferon-gamma) [20]

Moreover, genetic studies pointed out the importance of adhesion molecules in the susceptibility to the disease[21, 22]. In 2003, we reported that a rare haplotype of the *ICAM-1* gene (intercellular adhesion molecule 1, expressed at the surface of endothelial cells) was

under-transmitted to patients, suggesting a protective effect of this haplotype [21]. In 2009, a pathway meta-analysis combining two GWAS uncovered genetic variants that interacted in the CAM (Cell Adhesion Molecules) pathway potentially contributing to MS development[22]. This new strategy led to the identification of new genes that do not significantly contribute to MS susceptibility individually but play an important role in interaction with other genes of the biological pathway. An earlier protein interaction network-based pathway analysis (PINBPA) considered all the SNPs with a p-value below of 0.05, without any prior analysis on genes function. That study showed significant interactions among adhesion molecule-coding genes including *ICAM-1*, *ITGB2 (integrin beta 2)*, *ITGAM (integrin alpha M)*, *ITGA6 (integrin alpha 6)*, *CD58 (CD58 molecule)*, *CD2 (CD2 molecule)* and *CD4 (CD4 molecule)*. This observation, coming from a meta-analysis of two independent GWAS, suggests that the biological network of adhesion molecules is involved in MS susceptibility.

In order to decipher the role of the CAMs pathway in MS susceptibility, we conducted a candidate pathway analysis (CPA) focusing on adhesion molecules. We found 5 networks of the CAMs pathway enriched in low p-values for genes interacting synergistically to confer MS susceptibility.

Results

CAMs pathway selected genes

Using KEGG database and Sabiosciences website, we identified 76 genes of interest involved in the adhesion molecule pathway. More than 85% of them (66 out of 76) were involved in 3 sub-pathways: Adhesion, Adherens Junctions and Tight Junctions, and 10 encode for transcription factors highly involved in adhesion molecule genes regulation (see Supplementary Table 1).

Using data from 8 GWAS studies, we selected calculated p-values of identified CAMs genes (+/- 1kb from the 5' and 3' UTR). Seven GWAS (IMSGC UK, IMSGC US, ANZGene, GeneMSA DU, GeneMSA SW, GeneMSA US and BWH/MIGEN) constituted the D1 Dataset.[4] The 8th GWAS (IMSGC WTCCC2) constituted the D2 dataset [3] (See figure 1).

70 of the 76 genes were represented by at least one SNP in at least one of the 7 datasets included in D1, whereas 68 genes of the 76 were represented by at least one SNP in the WTCCC2 dataset (D2 dataset). 6 of the 76 genes of interest were thus excluded from further analysis. The number of available SNPs per gene and per dataset is given in Supplementary Table 2. D1 and D2 gene-wise p-values, computed using VEGAS software (see Materials and Methods section and Figure 1) are given in Table 1. As expected, *CD58*, *NFKB1* and *STAT3*

genes previously identified as associated with MS in GWAS show low gene-wise p-values, $P_{D1} = 1.89x10^{-6}$, $P_{D2} = 0$ (meaning that $P<89x10^{-6}$, see Materials and Methods section); $P_{D1} = 0.0125$, $P_{D2} = 0.0011$; $P_{D1} = 2.94x10^{-5}$, $P_{D2} = 1.5x10^{-5}$ respectively.

Interestingly, *VCAM1*, previously identified as associated to the disease, with rs11581062 being the most statistical significant SNP, did not reach a significant p-value (P_{D1} =0.7224; P_{D2} =0.993) neither in the D1 nor in the D2 datasets. The published associated SNP (rs11581062) is 202kb far from VCAM1 gene. We conclude that even if VCAM1 is a good candidate gene for MS physiopathology, not enough evidence support its involvement within the rs11581062 region. Furthermore, none CIS-e-QTL regulating VCAM1 expression was described within the 1 Mb region around the gene (Pritchard lab resources).

Sub-networks associated with MS susceptibility

As described by IMSGC et al. [23], we used a curated human protein interaction network (PIN) dataset, consisted in a network of more than 400,000 interactions among ~25,000 proteins. In order to be confident with the interactions, we selected those quoted in at least 2 publications. It results to a reduced human PIN dataset of 8,920 proteins and 27,724 interactions. Using Cytoscape software, we attributed gene-wise p-values to 70 of the 76 CAMs genes (listed in Supplementary Table 1). As mentioned in IMSGC et al. [23], Cytoscape plugin jActives modules was used to calculate a global score (Z-score) for all the possible networks that could be form from the PIN dataset, using D1 and D2 p-values. Networks with Z-scores greater than 3.0 are generally considered significant, i.e. these sub-networks are enriched in CAMs genes showing D1 and/or D2 gene-wise significant p-values. The cytoscape software was applied on the reduced human PIN dataset, containing CAMs genes (see supplementary Table 1 and Figure 2) and non-CAMs ones. 64 networks were highlighted regarding the enrichment of CAMs genes with significant p-Values. Focusing on the process of BBB transmigration by T-cells, we only considered as relevant for our study 6 networks with at least 50% of CAMs genes. Finally, we excluded subnets that were only enriched in CAMs transcription factors to prevent from non-specific association signal only driven by transcription factors (see Table 2 and Figure 2). We identified 5 networks constituted by genes known to interact together within the CAMs pathway and meeting all expressed criteria (See Figure 3).

Discussion

GWAS and replication studies have successfully identified approximately 110 non-MHC MS susceptibility genes[3-7]. However, an important part of genetic heritability remains to be discovered[3]. One of the hypotheses which can be put forward postulates that genes without individual effect could influence the susceptibility to the disease through genetic interactions [11]. Here, we report the first candidate pathway analysis on the Cell Adhesion Molecules (CAMs) network in MS. Our results highlight 5 sub-networks enriched in CAMs genes with significant p-values, reflecting a potential genetic contribution of these molecules in MS susceptibility.

In 2009, Baranzini et al. identified 7 genes of the CAMs pathway as associated with MS susceptibility considering their interaction[22]. Focusing on the CAMs pathway and using two powerful datasets (5,545+9,772 patients and 12,153+17,376 controls), our study refines the role of these genes as MS genetic factors.

Two of the 5 identified networks (network 2 and 27) appear to be the highest contributors to MS physiopathology. The network 2 is constituted by ICAM1, ICAM3 and ITGAL genes that are involved in the adhesion process of the T cells on the endothelial cells of the Blood-Brain Barrier (BBB). The adhesion process between endothelial and T cells is one of the earliest event in MS development leading to the T cells transmigration into the central nervous system across the BBB. Of note, ITGAL, ICAM1 and ICAM3 molecules constitute ligand-receptor complexes, ITGAL interacting both with ICAM1 and ICAM3. Our hypothesis on the role of these genes in MS susceptibility is that some intra- and inter-gene combinations of polymorphisms could influence the adhesion process of T cells on the BBB leading to a positive or negative modulation of the inflammatory cells flow into the CNS.

The network 27 is constituted by CD82 (KAI1), ITGA4, ITGB1, ITGB2, and HLA-DMA genes. The ITGB1 and ITGA4 genes code for the two subunits of the VLA4 molecule. CD82 is notably known for its inhibitory effect on ITGB1 activation[24]. The effective role of CD82 in inflammatory response has not been yet investigated. In the MS context, it would be relevant to analyze the development and the severity of EAE (Experimental Autoimmune Encephalomyelitis; MS animal model) in CD82 deficient mice. Since CD82 initiates the differentiation of oligodendrocytes precursors into mature myelinating cells[25], quantification of remyelination in this model would be of interest.

Interestingly we identified two networks containing the ITGA4 molecule which is the target of Natalizumab. The interaction between VLA-4 and his receptor VCAM1 is known to be crucial for the transmigration of T cells across the BBB and potentially the molecular basis for this treatment's effectiveness.

In conclusion, our results highlighting the role of CAMs genes interactions in MS susceptibility could be of high interest to identify new targets for efficient treatments for MS. A monoclonal antibody directed against ITGA4 that prevent its interaction with VCAM1 already exists[18]. We propose that targeting ITGAL or ICAM1 or ICAM3 in order to block their interactions could also be an efficient way to prevent the crossing of the BBB by T cells.

Materials and Methods

Ethics Statement

All the data used in this study have been generated by previous studies [3, 4]. Ethics approval could be found in the corresponding publications.

Selection of genes coding for CAMs pathway molecules

The selection of the genes belonging to the network has been performed in 3 phases: Using KEGG (Kyoto Encyclopedia for Genes and Genomes) database (http://www.genome.jp/kegg/) [26, 27], we searched for adhesion molecules involved in tight and adherens junctions between BBB endothelial cells, and for those involved in T cells transmigration. Still Using KEGG database, we then identified molecules interacting directly with adhesion molecules. In a last step, we selected the 10 transcription factors that are most involved in adhesion molecule-coding genes expression and regulation using the SabioSciences website (http://www.sabiosciences.com) through GeneCards database (http://www.genecards.org/) [28]. Using this 3-step strategy, we compiled a list of 76 genes of interest.

Datasets

Eight datasets were used in this study. Seven datasets (IMSGC UK, IMSGC US, ANZGene, GeneMSA DU, GeneMSA SW, GeneMSA US and BWH/MIGEN) are described in Patsopoulos et al and constitute the D1 dataset. [4]. The 8th dataset consists in the MS GWAS published by the IMSGC and WTCCC2 in 2011 and conbstitute the D2 dataset[3]. Since the IMSGC-WTCCC2 dataset is to date the most powerful GWAS published in MS, we decided to use this dataset separately to avoid signals losses.

Statistical analysis

Gene-wise p-values computation

With data from 8 GWAS we computed an individual gene-wise p-value corresponding to the association of each gene with MS in each GWAS using VEGAS software [29] (see Figure 1) VEGAS assigns SNPs to each of 17,787 autosomal genes according to positions on the UCSC Genome Browser (hg18 assembly). For the capture of regulatory regions and SNPs in LD, gene boundaries are defined as +/- 50kb of each gene. VEGAS takes into account LD patterns between markers within a gene by using Monte-Carlo simulations from the multivariate normal distribution on the basis of the LD structure of a set of reference individuals (the HapMap2 CEU [Utah residents with ancestry from northern and western Europe from the CEPH collection] population). In VEGAS, the number of simulations per gene is determined adaptively. In the first stage, 10^3 simulations are performed. If the resulting empirical p value is less than 0.1, 10^4 simulations are then performed. If the empirical p value from 10^4 simulations is less than 0.001, the program will perform 10^6 simulations. At each stage, the simulations are mutually exclusive. For computational reasons, if the empirical p value is 0, then no more simulations will be performed. An empirical p value of 0 from 10^6 simulations can be interpreted as $p < 10^{-6}$, which exceeds a Bonferroni-corrected threshold of $p < 2.8 \times 10^{-6}$.

To combine the gene-wise p-values across the seven datasets described above (named D1), we applied Fisher's method for each of the 70 selected genes. The gene-wise p-values from the IMSGC-WTCCC2 dataset were used separately, as a second independent dataset (named D2) (See Figure 1).

Protein interaction network-based pathway analysis (PINBPA)

As described by IMSGC et al. in 2013 [23], we integrated data from a curated human protein interaction network (PIN) dataset in Cytoscape software [30]. Cytoscape plugin jActives modules was used to calculate a global score for each network enriched in CAMs genes showing D1 and/or D2 gene-wise significant p-values. Using successive filters we identified networks enriched in low p-Values with at least 50% of CAMs genes and less than 50% of transcription factors (See Figure 2).

Acknowledgments.

This study was supported by the Institut National de la Santé et de la Recherche Médicale (INSERM), the Fondation pour la Recherche sur la Sclérose En Plaques (ARSEP), the Association Française contre les Myopathies (AFM), GIS-IBISA and ICM Carnot Institute. The research leading to these results has received funding from the program

"Investissements d'avenir" ANR-10-IAIHU-06. We thank ICM, CIC Pitié-Salpêtrière, Généthon, BRC-REFGENSEP's and IMSGC's members for their help and support as well as Jorge Oksenberg and Pierre-Antoine Gourraud. VD received a travel grant from the Fondation ARSEP and ICM Carnot Institute. Philip L. De Jager is a Harry Weaver Neuroscience Scholar of the National MS Society. SEB is a Harry Weaver Neuroscience fellow from the US National MS Society. This investigation was supported (in part) by a Postdoctoral Fellowship from the National Multiple Sclerosis Society to Nikolaos A. Patsopoulos (FG 1938-A-1).

The members of the **International Multiple Sclerosis Genetics Consortium (IMSGC)** are:
Maria Ban, Sergio Baranzini, Lisa Barcellos, Gary Beecham, Ashley Beecham, Luisa Bernardinelli, David Booth, Steffan Bos, Dorothea Buck, William Bush, Manuel Comabella, Alastair Compston, Chris Cotsapas, Isabelle Cournu-Rebeix, Bruce Cree, Sandra D'Alfonso, Mark Daly, Vincent Damotte, Mary Davis, Paul de Bakker, Philip L. De Jager, Benedicte Dubois, Federica Esposito, Bertrand Fontaine, An Goris, Pierre-Antoine Gourraud, Todd Green, Elisabeth Gulowsen Celius, Athena Hadjixenofontos, David Hafler, Jonathan Haines, Hanne F. Flinstad, Stephen Hauser, Clive Hawkins, Bernhard Hemmer, Jan Hillert, Rogier Hintzen, Dana Horáková, Adrian J. Ivinson, Anu Kemppinen, Jun-ichi Kira, Ingrid Kockum, Robin Lincoln, Roland Martin, Filippo Martinelli Boneschi, Jacob L. McCauley, Inger-Lise Mero, Jorge Oksenberg, Tomas Olsson, Annette Oturai, Aarno Palotie, Nikolaos Patsopoulos, Margaret Pericak-Vance, John Rioux, Janna Saarela, Stephen Sawcer, Nathalie Schnetz-Boutaud, Finn Sellebjerg, Helle Soendergaard, Per Soelberg Sorensen, Anne Spurkland, Jim Stankovich, Graeme Stewart, Bruce Taylor, Anna Ticca, Sandra West, Frauke Zipp

The members of the **Wellcome Trust Case Control Consortium 2 (WTCCC2)** are:

Management Committee

Peter Donnelly (Chair), Ines Barroso (Deputy Chair), Jenefer M. Blackwell, Elvira Bramon, Matthew A. Brown, Juan P. Casas, Aiden Corvin, Janusz Jankowski, Hugh S. Markus, Christopher G. Mathew, Colin N.A. Palmer, Robert Plomin, Anna Rautanen, Stephen Sawcer, Richard C. Trembath, Ananth C. Viswanathan, Nicholas W. Wood

Data and Analysis Group

Chris C.A. Spencer, Gavin Band, Céline Bellenguez, Colin Freeman, Garrett Hellenthal, Eleni Giannoulatou, Matti Pirinen, Richard Pearson, Amy Strange, Zhan Su, Damjan Vukcevic, Peter Donnelly

DNA, Genotyping, Data QC and Informatics Group

Cordelia Langford, Sarah E. Hunt, Sarah Edkins, Rhian Gwilliam, Hannah Blackburn, Suzannah J. Bumpstead, Serge Dronov, Matthew Gillman, Emma Gray, Naomi Hammond, Alagurevathi Jayakumar, Owen T. McCann, Jennifer Liddle, Simon C. Potter, Rathi Ravindrarajah, Michelle Ricketts, Matthew J Waller, Paul Weston, Sara Widaa, Pamela Whittaker, Ines Barroso, Panos Deloukas

Associate Members

Alexander Dilthey, Stephen Leslie, Loukas Moutsianas, Marc L. Perez, Gil McVean

Publications Committee

Christopher G. Mathew, Jenefer M. Blackwell, Matthew A. Brown, Aiden Corvin, Mark I McCarthy, Chris C.A. Spencer

Conflict of Interest.

Authors declare no conflict of interest

Supplementary Information.

Supplementary information is available at *Genes and Immunity*'s website.

References

1. Nylander A, Hafler DA. Multiple sclerosis. *The Journal of clinical investigation* 2012; **122**(4)**:** 1180-8.

2. Compston A, Coles A. Multiple sclerosis. *Lancet* 2008; **372**(9648)**:** 1502-17.

3. IMSGC, WTCCC2, Sawcer S, Hellenthal G, Pirinen M, Spencer CC *et al.* Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. *Nature* 2011; **476**(7359)**:** 214-9.

4. Patsopoulos NA, Esposito F, Reischl J, Lehr S, Bauer D, Heubach J *et al.* Genome-wide meta-analysis identifies novel multiple sclerosis susceptibility loci. *Ann Neurol* 2011; **70**(6)**:** 897-912.

5. Lill CM, Schjeide BM, Graetz C, Liu T, Damotte V, Akkad DA *et al.* Genome-wide significant association of ANKRD55 rs6859219 and multiple sclerosis risk. *Journal of medical genetics* 2013.

6. Lill CM, Schjeide BM, Graetz C, Ban M, Alcina A, Ortiz MA *et al.* MANBA, CXCR5, SOX8, RPS6KB1 and ZBTB46 are genetic risk loci for multiple sclerosis. *Brain* 2013; **136**(Pt 6)**:** 1778-82.

7. International Multiple Sclerosis Genetics C, Beecham AH, Patsopoulos NA, Xifara DK, Davis MF, Kemppinen A *et al.* Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis. *Nat Genet* 2013.

8. Bush WS, Sawcer SJ, de Jager PL, Oksenberg JR, McCauley JL, Pericak-Vance MA *et al.* Evidence for polygenic susceptibility to multiple sclerosis--the shape of things to come. *Am J Hum Genet* 2010; **86**(4)**:** 621-5.

9. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ *et al.* Finding the missing heritability of complex diseases. *Nature* 2009; **461**(7265)**:** 747-53.

10. Marian AJ. Elements of 'missing heritability'. *Current opinion in cardiology* 2012; **27**(3)**:** 197-201.

11. Wang K, Li M, Hakonarson H. Analysing biological pathways in genome-wide association studies. *Nat Rev Genet* 2010; **11**(12)**:** 843-54.

12. Ramanan VK, Shen L, Moore JH, Saykin AJ. Pathway analysis of genomic data: concepts, methods, and prospects for future development. *Trends in genetics : TIG* 2012; **28**(7)**:** 323-32.

13. Holman DW, Klein RS, Ransohoff RM. The blood-brain barrier, chemokines and multiple sclerosis. *Biochim Biophys Acta* 2011; **1812**(2)**:** 220-30.

14. Engelhardt B, Wolburg H. Mini-review: Transendothelial migration of leukocytes: through the front door or around the side of the house? *Eur J Immunol* 2004; **34**(11)**:** 2955-63.

15. Engelhardt B, Ransohoff RM. The ins and outs of T-lymphocyte trafficking to the CNS: anatomical sites and molecular mechanisms. *Trends Immunol* 2005; **26**(9)**:** 485-95.

16. Engelhardt B. Immune cell entry into the central nervous system: involvement of adhesion molecules and chemokines. *J Neurol Sci* 2008; **274**(1-2)**:** 23-6.

17. Man S, Ubogu EE, Ransohoff RM. Inflammatory cell migration into the central nervous system: a few new twists on an old tale. *Brain Pathol* 2007; **17**(2)**:** 243-50.

18. Yednock TA, Cannon C, Fritz LC, Sanchez-Madrid F, Steinman L, Karin N. Prevention of experimental autoimmune encephalomyelitis by antibodies against alpha 4 beta 1 integrin. *Nature* 1992; **356**(6364)**:** 63-6.

19. Polman CH, O'Connor PW, Havrdova E, Hutchinson M, Kappos L, Miller DH *et al.* A randomized, placebo-controlled trial of natalizumab for relapsing multiple sclerosis. *N Engl J Med* 2006; **354**(9)**:** 899-910.

20. Cayrol R, Wosik K, Berard JL, Dodelet-Devillers A, Ifergan I, Kebir H *et al.* Activated leukocyte cell adhesion molecule promotes leukocyte trafficking into the central nervous system. *Nat Immunol* 2008; **9**(2)**:** 137-45.

21. Cournu-Rebeix I, Genin E, Lesca G, Azoulay-Cayla A, Tubridy N, Noe E *et al.* Intercellular adhesion molecule-1: a protective haplotype against multiple sclerosis. *Genes Immun* 2003; **4**(7)**:** 518-23.

22. Baranzini SE, Galwey NW, Wang J, Khankhanian P, Lindberg R, Pelletier D *et al.* Pathway and network-based analysis of genome-wide association studies in multiple sclerosis. *Hum Mol Genet* 2009; **18**(11)**:** 2078-90.

23. IMSGC. Network-Based Multiple Sclerosis Pathway Analysis with GWAS Data from 15,000 Cases and 30,000 Controls. *Am J Hum Genet* 2013.

24. Lee HA, Park I, Byun HJ, Jeoung D, Kim YM, Lee H. Metastasis suppressor KAI1/CD82 attenuates the matrix adhesion of human prostate cancer cells by suppressing fibronectin expression and beta1 integrin activation. *Cellular physiology and biochemistry : international journal of experimental cellular physiology, biochemistry, and pharmacology* 2011; **27**(5)**:** 575-86.

25. Mela A, Goldman JE. CD82 blocks cMet activation and overcomes hepatocyte growth factor effects on oligodendrocyte precursor differentiation. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 2013; **33**(18)**:** 7952-60.

26. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic acids research* 2000; **28**(1)**:** 27-30.

27. Kanehisa M, Goto S, Sato Y, Furumichi M, Tanabe M. KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic acids research* 2012; **40**(Database issue)**:** D109-14.

28. Rebhan M, Chalifa-Caspi V, Prilusky J, Lancet D. GeneCards: integrating information about genes, proteins and diseases. *Trends in genetics : TIG* 1997; **13**(4)**:** 163.

29. Liu JZ, McRae AF, Nyholt DR, Medland SE, Wray NR, Brown KM *et al.* A versatile gene-based test for genome-wide association studies. *Am J Hum Genet* 2010; **87**(1)**:** 139-45.

30. Cline MS, Smoot M, Cerami E, Kuchinsky A, Landys N, Workman C *et al.* Integration of biological networks and gene expression data using Cytoscape. *Nature protocols* 2007; **2**(10)**:** 2366-82.

Legend

Figure 1. D1 and D2 P-Values computation
Colored circle represent SNPs. The number of SNPs varies for the different genes (from 1 to 539, see supplementary table 2) in the different datasets.
Data from each GWAS are composed by p-value assigned to each SNPs localized within 70 CAMs genes (gene boundaries +/- 50Kb).
Dataset-A: GWAS ANZGene; Dataset-B: GWAS BWH/MIGEN; Dataset-C: GWAS MSA DU; Dataset-D: GWAS MSA SW; Dataset-E: GWAS MSA US; Dataset-F: GWAS IMSGC US; Dataset-G: GWAS IMSGC UK; Dataset-H: GWAS WTCCC2.
VEGAS Software assigns one gene-wise p-Value to each gene in each GWAS
Fischer product assigns D1 p-Value to each gene.

Figure 2. Protein interaction network-based pathway analysis

Figure 3. Networks enriched in CAMs genes with significant D1 and/or D2 pvalue
Each node represents a gene product and each edge, a physical interaction reported in at least two independent publications. Each network is enriched with CAMs genes having low p-values. Nodes are colored according to D1 p-values (red: highly significant, orange: moderate, yellow: D1_p-value = 0.05, white: not significant). Analysis was performed using Cytoscape and *jActive module plugin*.

Table 1. D1_ and D2_ Gene-wise p-values.
D1_Gene-wise p-values have been computed using VEGAS software and fisher product. D2_Gene-wise p-values were calculated using VEGAS software only.

Table 2. Selection criteria of the sub-networks
In red, the 5 selected sub-networks that met the criteria.
In white, excluded sub-networks with less than 50% of CAMs genes in the sub-network
In blue, excluded sub-networks with more than 50% of CAMs transcription factors among the CAMs genes in the sub-network
Genes: Number of genes in the sub-network
CAMs Genes: Number of genes of the network that belong to the CAMs genes list
% CAMs Genes: Ratio (# Genes / # CAMs Genes)
# CAMs TF: Number of CAMs Transcription factors in the sub-network
CAMs TF / CAMs Genes: Ratio (# CAMs TF / # CAMs Genes)

Supplementary Table 1. Name, localization and function of selected CAMs pathway genes.
Chr: chromosome
Start: Starting Chromosomal Position in NCBI B36
End: Ending Chromosomal Position in NCBI B36
TJ: Tight Junctions

AJ: Adherens Junctions
TF: Transcription Factors

Supplementary Table 2. Number of genotyped SNPs in each original dataset.